AD-A099 357 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
MORSE PROGRAMS: A TOPOLOGICAL APPROACH TO SMOOTH CONSTRAINED OP--ETC(U)
JAN 81 O FUJIWARA DAAG29-80-C-0041
UNCLASSIFIED MRC-TSR-2169 NL

1 OF 1
AD
A099357

END
DATE
FILMED
6-81
DTIC

AD A099357

MRC Technical Summary Report #2169

MORSE PROGRAMS: A TOPOLOGICAL APPROACH TO SMOOTH CONSTRAINED OPTIMIZATION

Okitsugu Fujiwara

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

Janu~~ary~~ 1981

(Received December 16, 1980)

MAY 27 1981

A

DTIC FILE COPY

Approved for public release
Distribution unlimited

Sponsored by

U.S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

and

National Science Foundation
Washington, D.C. 20550

81

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

MORSE PROGRAMS: A TOPOLOGICAL APPROACH TO SMOOTH CONSTRAINED OPTIMIZATION

Okitsugu Fujiwara

Technical Summary Report # 2169

January 1981

ABSTRACT

We consider nonlinear constrained optimization problems whose objective and constraint functions are sufficiently smooth. No convexity is assumed.

Our basic tools are from differential topology. We show that these problems can be reduced to the study of minimizing a Morse function on a manifold with boundary and we give the geometrical meaning to the first order conditions, the second order sufficiency conditions, and strict complementary slackness condition.

Our main concerns are the second order sufficiency conditions, sensitivity analysis, generic properties of smooth nonlinear programs, global duality, local uniquenss, and strict complementary slackness.

AMS (MOS) Subject Classifications: 58A05, 90C30

Key Words: Morse functions, Nonlinear programming, Second order sufficiency conditions, Strict complementary slackness, Sensitivity analysis

Work Unit Number 5: Operations Research

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. This is the revised version of Cowles Foundation Discussion Paper No. 539. (Yale University) supported in part by National Science Foundation Grants ENG-78-25182 and SOC-77-03277.

SIGNIFICANCE AND EXPLANATION

Nonlinear optimization problems arise in economic theory, in management science and in other fields. In the analysis of global optima of such problems, we quite often assume the functions concerned are convex. But in general those functions cannot be expected to be convex.

In this paper it is assumed that those functions are not necessarily convex but sufficiently smooth. We show that almost always nonlinear optimization problems have a unique global solution if global solutions exist, and we also show that with slightly perturbed data of a special type, those global optima almost always change smoothly in a certain problem.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

MORSE PROGRAMS: A TOPOLOGICAL APPROACH TO SMOOTH CONSTRAINED OPTIMIZATION

Okitsugu Fujiwara

1. Introduction

The nonlinear programming problem

$$(Q): \qquad \text{minimize}\{f(x) \text{ subject to } g(x) \leq b\}$$

where $x \in R^n$, $b \in R^m$ is called a convex program if f and g are convex. Convex programs enjoy a number of desirable global properties (e.g. Mangasarian [12], Rockafellar [13]) which do not hold in nonconvex programs. But these properties are true locally under certain constraint qualifications (e.g. Fiacco and McCormick [6], Avriel [2]). An important question is: do these constraint qualifications hold for almost all nonlinear programs? This question was recently answered affirmatively by Spingarn and Rockafellar [17] who showed, assuming differentiability of the objective and constraint functions, at any local minimum point x^* of $(Q(u,v))$, where

$$(Q(u,v)): \quad \text{minimize}\{f(x) - u^T x \text{ subject to } g(x) \leq b + v\},$$

that the Jacobian matrix of g at x^* has full rank; the strict complementary slackness condition; and the second order sufficiency conditions hold at x^*, for almost every (u,v) in $R^n \times R^m$.

However their clever argument is analytic and devoid of geometrical intuition. Spingarn ([14], [15], [16]) has provided a geometrical interpretation of his results using his notion "cyrtohedra", a generalization of manifolds with corners.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. This is the revised version of Cowles Foundation Discussion Paper No. 539 (Yale University) supported in part by National Science Foundation Grants ENG-78-25182 and SOC-77-03277.

The purpose of this paper is also to give a geometrical answer to the question: do the strong second order sufficiency conditions hold at any local minimum point for almost all nonlinear programs? Our idea is to reduce the nonlinear programming problem to a finite family of "well-behaved" nonlinear programs by perturbing the objective function in a linear fashion and perturbing the right hand side of the constraints by adding a constant. Each of the "well-behaved" nonlinear programs will consist of minimizing a Morse function on a manifold with boundary, where the Morse function has no critical points on the boundary. The constraint set being a manifold with boundary is the geometrical meaning of the full rank condition of the Jacobian; the objective function being a Morse function is the geometrical meaning of the second order sufficiency conditions; the lack of critical points on the boundary is the geometrical meaning of strict complementary slackness condition. Moreover, our perturbation gives us a unique global solution.

We follow a classical tradition of first studying an equality constrained program, in which the feasible region is a manifold without boundary; and then reducing an inequality constrained program to a finite family of constrained programs whose constraints consist of a finite set of equalities and one inequality (through the device of active or binding constraints), where we decompose the feasible region into a finite number of manifolds with boundary.

Our main concerns are the second order sufficiency conditions (Theorems A, F); sensitivity analysis (Theorems B, E); generic properties of smooth nonlinear programs (Theorems C, H); strict complementary slackness condition (Theorem G), and local uniqueness (Theorem E).

2. Basic Definitions and Notation

A property that holds except on a subset of R^n whose Lebesgue measure is zero is said to hold at <u>almost every</u> $u \in R^n$. The complement of a measure zero set in R^n is said to have <u>full measure</u> in R^n.

The Jacobian matrix and the Hessian matrix of f at x are denoted by $Df(x)$ and $D^2f(x)$ respectively.

Let $f : M \to R^m$ be a C^γ map from a k-dimensional C^γ manifold M with <u>boundary</u> ∂M in R^n. Let (ϕ, U) be a <u>local parametrization</u> of M at x such that $x = \phi(u)$, $u \in U \subseteq H^k = \{x \in R^k | x_k \geq 0\}$. The <u>tangent space</u> T_xM of M at x is defined to be the image of $D\phi(u) : R^k \to R^n$. A point $x \in M$ is a <u>regular point</u> of f if $D(f\phi)(u) : R^k \to R^m$ is surjective, otherwise x is a <u>critical point</u> of f. A critical point x of $f : M \to R^1$ is <u>nondegenerate</u> if the $k \times k$ matrix $D^2(f\phi)(u)$ is nonsingular. It is easily shown that the above definitions do not depend on the choice of local parametrization. A point $y \in R^m$ is a <u>regular value</u> of f, denoted by $f \pitchfork y$, if every $x \in f^{-1}(y)$ is a regular point of f, otherwise y is a <u>critical value</u> of f. $f : M \to R^1$ is a <u>Morse function</u> if all critical points of f are nondegenerate.

Let $f : M \to N$ be a C^γ map, $A \subseteq N$ be a C^γ submanifold of N. f is <u>transversal</u> to A, denoted by $f \pitchfork A$, if for every $x \in f^{-1}(A)$, Image $Df(x) + T_{f(x)}A = T_{f(x)}N$ holds, where $Df(x) : T_xM \to T_{f(x)}N$ is the derivative of f. Two submanifolds A, B of M are <u>transversal</u> denoted by $A \pitchfork B$, if $i \pitchfork B$ where $i : A \to M$ is the inclusion map. f is an <u>immersion</u> if for every $x \in M$, $Df(x) : T_xM \to T_{f(x)}N$ is injective. f is a <u>submersion</u> if $Df(x)$ is surjective for every $x \in M$. f is <u>proper</u> if the preimage of every compact set in N is compact in M. An immersion that is injective and proper is called an <u>embedding</u>.

We refer the interested reader to Guillemin and Pollack [8] for an introduction to the concepts of differential topology that will be used in this paper. Those theorems of elementary differential topology which are used in the body of this paper are stated in the appendix. The proofs of those theorems can be found in Gillemin/Pollack [8] and Hirsch [10].

3. Equality Constraints: Properties of Morse Programs

Throughout this section we consider a program

$$(P): \qquad \text{minimize } \{f(x) \text{ subject to } g(x) = b\}$$

and a perturbation of (P)

$$(P(u,v)): \quad \text{minimize } \{f(x) - u^T x \text{ subject to } g(x) = b + v\}$$

where $f : R^n \to R^1$, $g : R^n \to R^m$; $u \in R^n$, $v \in R^m$; $n > m$, and we assume f and g are of class C^2.

Definition. A program (P) is a Morse program if $g \pitchfork b$ and f is a Morse function on $g^{-1}(b)$.

Definition. A point $x \in g^{-1}(b)$ is a critical point of (P) if x is a critical point of f on $g^{-1}(b)$.

It is easily verified that nondegenerate critical points are isolated (cf. Guillemin/Pollack [8]). Hence each critical point of a Morse program (P) is isolated. By the Morse Lemma (Appendix (1)) the local behavior of a function at a nondegenerate critical point is completely determined, i.e., at any critical point of a Morse program (P), f has a strict local minimum, a strict local maximum, or a saddle point.

If $g \pitchfork b$ and $g \in C^\gamma$ then $g^{-1}(b)$ is $(n-m)$-dimensional C^γ submanifold of R^n (Appendix (5)).

A Morse program has three distinguishing properties:

(a) The second order sufficiency conditions hold at every local minimum point of a Morse program (P) (Theorem A).

(b) If x is a critical point of a Morse program (P), then the associated Lagrange multiplier λ exists and

the matrix $\begin{bmatrix} D^2f(x) + \sum_{i=1}^{m} \lambda_i D^2 g_i(x) & Dg(x)^T \\ Dg(x) & 0 \end{bmatrix}$ is non-singular (Theorem B).

(c) Generically (P) can be considered a Morse program, namely $(P(u,v))$ is a Morse program for almost every $(u,v) \in R^n \times R^m$ (Theorem C).

We will discuss the existence of the Lagrange multiplier and its uniqueness geometrically, without using Farkas lemma.

Suppose $g \pitchfork b$ and $g \in C^\gamma$ $(\gamma \geq 2)$. Then $M = g^{-1}(b)$ is $(n-m)$-dimensional C^γ submanifold of R^n and at each point $x \in M$ $Dg(x)$ has full rank, hence $R^n = \operatorname{Ker} Dg(x) \oplus \operatorname{Im} Dg(x)^T$ and $T_xM = \operatorname{Ker} Dg(x)$ (because differentiating $g\phi \equiv b$ on U, where (ϕ, U) is a local parametrization of M at $x = \phi(p)$, *we obtain* $T_xM = \operatorname{Im} D\phi(p) \subseteq \operatorname{Ker} Dg(x)$. Comparing dimensions of both sides we have $T_xM = \operatorname{Ker} Dg(x)$). A point $x \in M$ is a critical point of f on M iff $Df(x)^T \perp T_xM$, because $D(f\phi)(p)R^{n-m} = Df(x)D\phi(p)R^{n-m}$ $= Df(x)T_xM = \{0\}$ iff $Df(x)^T \perp T_xM$. Then $Df(x)^T \in \operatorname{Im} Dg(x)^T$.

Hence we have

<u>Lemma 1.</u>[1)] <u>If</u> $g \pitchfork b$, <u>then</u> $x \in g^{-1}(b)$ <u>has a Lagrange multiplier iff</u> x <u>is a critical point of</u> f <u>on</u> $g^{-1}(b)$. <u>Moreover the Lagrange multiplier is uniquely determined.</u>

The next lemma gives a representation of the Hessian matrix of f at $x \in g^{-1}(b)$, in terms of the seond derivative of the Lagrangian at x.

---

1) This fact has been pointed out previously by Tanabe ([18] Proposition 1).

Lemma 2.[1] Let $g \pitchfork b$, x be a critical point of $M = g^{-1}(b)$ with the associated Lagrange multiplier λ, $L(x) = D^2 f(x) + \sum_{i=1}^{m} \lambda_i D^2 g_i(x)$ and (ϕ, U) be a local parametrization of M at x such that $x = \phi(p)$ for $p \in U \subseteq R^{n-m}$. Then $D^2(f\phi)(u) = D\phi(p)^T L(x) D\phi(p)$.

Proof.[2] By the chain rule we have

$$D^2(f\phi)(p) = D\phi(p)^T D^2 f(x) D\phi(p) + \sum_{j=1}^{n} \frac{\partial f(x)}{\partial x_j} D^2\phi_j(p) \quad . \tag{3.1}$$

Differentiating $\sum_{i=1}^{m} \lambda_i (g_i \phi) \equiv \sum_{i=1}^{m} \lambda_i b_i$ on U, we have

$$D\phi(p)^T \Big(\sum_{i=1}^{m} \lambda_i D^2 g_i(x) \Big) D\phi(p) + \sum_{j=1}^{m} \left[\sum_{i=1}^{m} \lambda_i \frac{\partial g_i(x)}{\partial x_j} \right] D^2\phi_j(p) = 0 \quad . \tag{3.2}$$

Adding (3.2) to (3.1) and taking account $Df(x) + \sum \lambda_i Dg_i(x) = 0$, we obtain $D^2(f\phi)(p) = D\phi(p)^T L(x) D\phi(p)$.

Q.E.D.

For $s \in T_x M$, $L(x)s$ is in R^n but not necessarily in $T_x M$. To obtain a linear homomorphism on $T_x M$, we project $L(x)s$ orthogonally onto $T_x M$. We denote this linear homomorphism on $T_x M$ by $L_M(x)$, which we call the induced homomorphism of $L(x)$ on $T_x M$ (Luenberger [11], 10.4). Let (ϕ, U) be a local parametrization of M at x such that $x = \phi(p)$, $p \in U \subseteq R^{n-m}$. We can choose (ϕ, U) so that the column vectors of $D\phi(p)$ are orthonormal in R^n. Then it is easily shown that the matrix representation of $L_M(x)$ with respect

---

1) This fact has been pointed out previously by Tanabe ([19] Lemma 5.4).

2) The idea for this proof was first given by Luenberger [11], 10.3.

to the column vectors of $D\phi(p)$, which is an orthonormal basis of T_xM, is $D\phi(p)^T L(x) D\phi(p)$ (Luenberger [11], 10.4). Hence by Lemma 2 we obtain

<u>Lemma 3</u>. <u>Let</u> $g \pitchfork b$ <u>and let</u> x <u>be a critcal point of</u> f <u>on</u> $M = g^{-1}(b)$. <u>Then</u>

$$x \ \underline{\text{is nondegenerate iff}}\ L_M(x)\ \underline{\text{is an isomorphiam}}\ .$$

Note that if x is nondegenerate, then $L(x)|_{T_xM}$ is $1 - 1$ since $L(x)D\phi(p)$ is $1 - 1$, and we have

$$L(x)T_xM \cap \text{Ker } D\phi(p)^T = \{0\} \ . \tag{3.3}$$

If, on the other hand, $L(x)T_xM \cap \text{Ker } D\phi(p)^T \neq \{0\}$, then $\dim\{L(x)T_xM \cap \text{Ker } D\phi(p)^T\} \geq 1$. Hence we have

$$\begin{aligned}\dim \text{Im}\{D\phi(p)^T L(x) D\phi(p)\} &= \dim D\phi(p)^T L(x)T_xM \\ &= \dim L(x)T_xM - \dim\{L(x)T_xM \cap \text{Ker } D\phi(p)^T\} \\ &< n - m\end{aligned}$$

which contradicts the nondegeneracy of x.

Lemma 3 shows that a Morse function, whose critical points are all nondegenerate, is an appropriate concept for the analysis of the second order optimality conditions. Summarizing the above argument, we obtain the first property of a Morse program.

<u>Theorem A</u>. <u>Let</u> (P) <u>be a Morse program and</u> x <u>be a critical point of</u> (P). <u>Then we have</u>

(a) $Dg(x)$ <u>has full rank</u>

(b) <u>there exists a unique</u> $\lambda \in R^m$ <u>such that</u> $Df(x)^T + Dg(x)^T\lambda = 0$

(c) $L(x) = D^2 f(x) + \sum_{i=1}^{m} \lambda_i D^2 g_i(x)$ <u>induces an isomorphism on</u> T_xM <u>where</u> $M = g^{-1}(b)$.

(d) <u>on</u> T_xM, $L(x)$ <u>is positive definite iff</u> x <u>is a local minimum; negative definite iff</u> x <u>is a local maximum; indefinite iff</u> x <u>is a saddle point</u>.

<u>Proof</u>. (a), (b), and (c) follow from, respectively, $g \pitchfork b$, Lemma 1, and Lemma 3. (d): positive (negative) definite $\Rightarrow$ local minimum (maximum) is obvious. If x is a local minimum (maximum), then $L(x)$ is positive (negative) semidefinite on T_xM. However, since $s^T L(x)s = s^T L_M(x)s$ for $s \in T_xM$, by Lemma 3 $L(x)$ must be positive (negative) definite on T_xM. The saddle point case is an immediate consequence of the preceding argument.

Q.E.D.

Now let us vary the right hand side $b \in R^m$ and consider a critical point x of (P) as a function of b, denoted by $x(b)$. A sufficient condition that $x(\cdot)$ is a C^1 function of b is the nonsingularity of the matrix

$$\begin{pmatrix} L(x) & Dg(x)^T \\ Dg(x) & 0 \end{pmatrix} \tag{3.4}$$

(this follows from the implicit function theorem).

Consider the function $F_b : R^n \times R^m \to R^n \times R^m$ defined by $F_b(x,\lambda) := (Df(x)^T + Dg(x)^T\lambda,\ g(x) - b)$. Then the nonsingularity of (3.4) for every critical point x and its associated Lagrange multiplier λ, is equivalent to $F_b \pitchfork 0$, which is equivalent to f being a Morse function on $M = g^{-1}(b)$; namely we have

<u>Theorem B</u>. <u>Let</u> $g \pitchfork b$. <u>Then</u>

$$F_b \pitchfork 0 \ \ \underline{\text{iff}}\ \ f\ \ \underline{\text{is a Morse function on}}\ \ M = g^{-1}(b) \quad .$$

Proof.

Let (ϕ, U) be a local parametrization of M at x such that $x = \phi(p)$ for $p \in U \subseteq R^{n-m}$.

(If) Let $(x,\lambda) \in F_b^{-1}(0)$, then x is a critical point of f on M (Lemma 1) and x is nondegenerate because f is a Morse function on M. Suppose $\begin{pmatrix} L(x) & Dg(x)^T \\ Dg(x) & 0 \end{pmatrix}\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} L(x)s + Dg(x)^T t \\ Dg(x)s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Then $s \in \text{Ker } Dg(x) = T_xM$ and $L(x)s = -Dg(x)^T t \in \text{Im } Dg(x)^T$. Note that $\text{Im } Dg(x)^T = \text{Ker } D\phi(p)^T$. Because we have $\text{Im } Dg(x)^T = (\text{Ker } Dg(x))^{\perp}$ (orthogonal complement of $\text{Ker } Dg(x)$ in R^n), $\text{Ker } D\phi(p)^T = (\text{Im } D\phi(p))^{\perp}$, and $\text{Ker } Dg(x) = T_xM = \text{Im } D\phi(p)$. Hence $L(x)s \in L(x)\, T_xM \cap \text{Ker } D\phi(p)^T$, so by (3.3) $L(x)s = 0$. Since $L(x)$ is $1-1$ on T_xM, this implies $s = 0$. Hence $t = 0$ since $Dg(x)^T$ is $1-1$. Therefore, we obtain $\text{Ker } DF_b(x,\lambda) = \{0\}$ for any $(x,\lambda) \in F_b^{-1}(0)$. Hence $F_b \pitchfork 0$.

(Only if)

Let x be a critical point of f on M. Then there exists $\lambda \in R^m$ such that $F_b(x,\lambda) = 0$ by Lemma 1. Suppose $D\phi(p)^T L(x) D\phi(p) r = 0$ for some $r \in R^{n-m}$. Let $s = D\phi(p)r$, then $L(x)s \in \text{Ker } D\phi(p)^T = \text{Im } Dg(x)^T$ hence $L(x)s = Dg(x)^T t$ for some $t \in R^m$. Then $\begin{pmatrix} L(x) & Dg(x)^T \\ Dg(x) & 0 \end{pmatrix}\begin{pmatrix} s \\ -t \end{pmatrix} = \begin{pmatrix} L(x)s - Dg(x)^T t \\ Dg(x)s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ because $s \in \text{Im } D\phi(p) = T_xM = \text{Ker } Dg(x)$. Hence $s = 0$ and $r = 0$ because $DF_b(x,\lambda)$ is nonsingular and $D\phi(p)$ is $1-1$. Therefore $D\phi(p)^T L(x) D\phi(p)$ is $1-1$, hence nonsingular, and by Lemma 3 x is nondegenerate.

Q.E.D.

The third property of Morse programs is genericity. In general (P) is not necessarily a Morse program, but we have,

Theorem C.[1] *If* $f \in C^2$ *and* $g \in C^{n-m+1}$, *then for almost every fixed* $v \in R^m$, $(p(u,v))$ *is a Morse program having at most one global solution for almost every* $u \in R^n$.

Proof. By Sard's Theorem (Appendix (2)) if $g : R^n \to R^m$ is of class C^{n-m+1}, then $g \pitchfork b+v$ for almost every $v \in R^m$. For a C^2 manifold $X \subseteq R^n$ and a C^2 map $h : X \to R^1$, $h(x) - u^T x$ is a Morse function for almost every $u \in R^n$ (Appendix (6)). Therefore for $v \in R^m$ such that $g \pitchfork b+v$, $f(x) - u^T x$ is a Morse function on $g^{-1}(b+v)$ for almost every $u \in R^n$. By Araujo and Mas-Colell ([1], Theorem 1),[2] we have

> Fix any $v \in R^m$, then for almost every $u \in R^n$
> $(P(u,v))$ has at most one global solution.

Q.E.D.

---

1) It can be shown that

> *If* $f \in C^2$ *and* $g \in C^{n-m+1}$, *then for almost every* $(u,v) \in R^n \times R^m$, $(P(u,v))$ *is a Morse program*.

(See the remark previous to Theorem E in section 4.)

2) Truman Bewley suggested the use of the Araujo/Mas-Colell theorem. For our application, their theorem can be stated "*Let* X *be a subset of* R^n, $\varphi : X \to R^1$ *be continuous, and* $\Phi : X \times R^n \to R^1$ *be defined by* $\Phi(x,u) = \varphi(x) - u^T x$ *for* $x \in X$, $u \in R^n$. *Then the function* $\Phi(\cdot,u) : X \to R^1$ *has at most one minimizer for almost every* $u \in R^n$." For our application for Theorem C, let $X = g^{-1}(b+v)$ and $\Phi(x,u) = f(x) - u^T x$.

4. Equality Constraints: Global Properties of Proper Morse Programs

A mapping $g : R^n \to R^m$ is called proper if the preimage of every compact set in R^m is compact in R^n. It is easily shown that g is proper if and only if

$$\{x_k\} \subseteq R^n, \ \|x_k\| \to \infty \Rightarrow \|g(x_k)\| \to \infty$$

where $\|\cdot\|$ is the Euclidean norm.

Definition. (Brown, Heal and Westhoff [3])

A program (P) is called proper if g is proper.

In this section we consider some global properties of proper Morse programs - global duality (Theorem D) and local uniqueness (Proposition 6, Theorem E).

A proper program has at least one global solution since $g^{-1}(b)$ is compact, hence by Araujo/Mas-Colell [1] if (P) is proper (P(u,0)) has a unique global solution for almost every $u \in R^n$. We will consider a family of parametrized programs

$$(P(y)): \qquad \underset{x \in R^n}{\text{minimize}} \{f(x) \ \text{ subject to } \ g(x) = y\} \ ,$$

and its global optimum value function

$$\omega(y) := \text{minimum}\{f(x) \ \text{ subject to } \ g(x) = y\} \ .$$

We also consider a program

$$(P_K): \qquad \underset{x \in K}{\text{minimize}} \{f(x) \ \text{ subject to } \ g(x) = b\} \ ,$$

and its dual

(D): $$\underset{\lambda \varepsilon R^m}{\text{maximize}}\ \phi_\sigma(\lambda) \quad ,$$

where

$$\phi_\sigma(\lambda) = \underset{x \varepsilon K}{\text{minimum}}\ \{f(x) + \lambda^T(g(x) - b) + \frac{\sigma}{2}\|g(x) - b\|^2\} \quad ,$$

K is a compact set of R^n, and $\sigma > 0$.

Since K is compact, there exists a global minimizer of $\phi_\sigma(\lambda)$ for any $\lambda \varepsilon R^m$ and $\sigma > 0$.

Hestenes showed

<u>Theorem</u> ([9] Chapter 5, Theorem 4.4)

<u>If</u> x^* <u>is a unique global minimum of</u> (P_K) <u>such that</u> $Df(x^*)^T + Dg(x^*)^T\lambda^* = 0$ <u>for some</u> $\lambda^* \varepsilon R^m$ <u>and</u> $D^2f(x^*) + \sum_1^m \lambda_i^* D^2 g_i(x^*)$ <u>is positive definite on</u> Ker $Dg(x^*)$, <u>then there exists</u> $\sigma_0 > 0$ <u>such that for any</u> $\sigma > \sigma_0$, x^* <u>is a unique global solution of</u> $\phi_\sigma(\lambda^*)$ <u>and hence</u> $\phi_\sigma(\lambda^*) = f(x^*)$.

As a matter of fact, we can claim

$$\phi_\sigma(\lambda^*) = \max_\lambda \phi_\sigma(\lambda)$$

namely we have

<u>Theorem D</u>

<u>If</u> g <u>is a proper Morse program having a unique global solution</u> x^* <u>with the associated Lagrange multiplier</u> λ^*, <u>and if we take</u> $K \supseteq g^{-1}(b)$,

then there exists $\sigma_0 > 0$ *such that for any* $\sigma > \sigma_0$, x^* *is a unique global solution of* $\phi_\sigma(\lambda^*)$ *and*

$$\phi_\sigma(\lambda^*) = \max_\lambda \phi_\sigma(\lambda) = \omega(b) = f(x^*) \quad .$$

Remark

The assumption is satisfied almost always if $f \in C^2$ and $g \in C^{n-m+1}$.

Proof

Since $K \supseteq g^{-1}(b)$, we have $\omega(b) = f(x^*)$, hence it suffices to show $\phi_\sigma(\lambda^*) = \max_\lambda \phi_\sigma(\lambda)$, which follows from (4.1) in the next lemma.

Q.E.D.

Lemma 4

(a) $\phi_\sigma(\cdot)$ *is a concave function of* λ *for any* $\sigma > 0$

(b) *For any* $\lambda \in R^m$, $g(x_\lambda) - b$ *is a supergradient*[1] *of* ϕ_σ *at* λ, *where* x_λ *is a global minimizer of* $\phi_\sigma(\lambda)$.

Proof

(a) is trivial. (b): We will show that for any $\mu \in R^m$,

$$\phi_\sigma(\mu) \leq \phi_\sigma(\lambda) + (\mu-\lambda)^T(g(x_\lambda) - b) \quad . \tag{4.1}$$

By the definition of x_λ, we have

$$\phi_\sigma(\mu) = f(x_\mu) + \mu^T(g(x_\mu) - b) + \frac{\sigma}{2} \|g(x_\mu) - b\|^2 \quad , \tag{4.2}$$

$$\phi_\sigma(\lambda) = f(x_\lambda) + \lambda^T(g(x_\lambda) - b) + \frac{\sigma}{2} \|g(x_\lambda) - b\|^2 \quad , \tag{4.3}$$

---

1) Rockafellar ([13] §23)

and

$$(4.4) \qquad \phi_\sigma(\mu) \le f(x_\lambda) + \mu^T(g(x_\lambda) - b) + \frac{\sigma}{2}\|g(x_\lambda) - b\|^2 \quad .$$

Substituting (4.2) and (4.3) into (4.4), we obtain (4.1).

Q.E.D.

Let us define a function $F : R^n \times R^m \times R^m \to R^n \times R^m$ by

$$F(x,\lambda,y) := (Df(x)^T + Dg(x)^T\lambda,\ g(x) - y) \quad .$$

We define $F_y(x,\lambda)$ as $F(x,\lambda,y)$. The next lemma is a key step toward our sensitivity analysis of proper Morse programs.

Lemma 5 (cf. Brown/Heal/Westhoff [3])

If g *is proper, then*

(a) $(P(y))$ *is a Morse program*

$$\Longleftrightarrow y \in Y := \{y \in R^m \mid g \pitchfork y,\ F_y \pitchfork 0\}$$

(b) Y *is open in* R^m

Proof (a) is equivalent to Theorem B.

(b): We claim that $\{y|g \pitchfork y\}$ is open in R^m and $\{y|g \pitchfork y,\ F_y \pitchfork 0\}$ is open in $\{y|g \pitchfork y\}$. Since both proofs are similar, we omit the first proof. So we will prove $\{y|g \pitchfork y,\ F_y \pitchfork 0\}$ is open in $\{y|g \pitchfork y\}$ which is open in R^m.

Suppose it is not open at $y^0 \in \{y|g \pitchfork y,\ F_y \pitchfork 0\}$. Then there exist $y^n \in \{y|g \pitchfork y\}$, $(x^n, \lambda^n) \in F_{y^n}^{-1}(0)$ such that $y^n \to y^0$ and $DF_{y^n}(x^n, \lambda^n)$ is

singular. Let K be a compact neighborhood of y^0 such that $K \subseteq \{y | g \pitchfork y\}$ and hence $Dg(x)$ has full rank for any $x \in g^{-1}(K)$.

Now for sufficiently large n, $x^n \in g^{-1}(K)$. Since g is proper and $g^{-1}(K)$ is compact, there exists a subsequence $\{x^{n_j}\}$ of $\{x^n\}$ such that $x^{n_j} \in g^{-1}(K)$, $x^{n_j} \to x^0$ for some $x^0 \in g^{-1}(K)$. Since $(x^{n_j}, y^{n_j}) \to (x^0, y^0)$ and $g(x^{n_j}) = y^{n_j}$, we have $g(x^0) = y^0$. By $g \pitchfork y^n$ we have

$$\text{(4.5)} \qquad \lambda^n = \lambda(x^n) := -\left(Dg(x^n)Dg(x^n)^T\right)^{-1} Dg(x^n)Df(x^n)^T \quad .$$

Since $\lambda(\cdot)$ is a continuous function of x on $g^{-1}(K)$, and since x^{n_j}, $x^0 \in K$ and $x^{n_j} \to x^0$, we have $\lambda^{n_j} \to \lambda^0 := \lambda(x^0)$. Then we obtain

$$(x^{n_j}, \lambda^{n_j}, y^{n_j}) \to (x^0, \lambda^0, y^0) \quad \text{and} \quad 0 = F_{y^n}(x^{n_j}, \lambda^{n_j}) \to F_{y^0}(x^0, \lambda^0) = 0 \quad .$$

By $F_{y^0} \pitchfork 0$, $DF_{y^0}(x^0, \lambda^0)$ is nonsingular. However we have also $DF_{y^{n_j}}(x^{n_j}, \lambda^{n_j}) \to DF_{y^0}(x^0, \lambda^0)$, hence $DF_{y^{n_j}}(x^{n_j}, \lambda^{n_j})$ is nonsingular for sufficiently large n_j, which contradicts our assumption. Therefore $\{y | g \pitchfork y, F_y \pitchfork 0\}$ is open in $\{y | g \pitchfork y\}$.

Q.E.D.

<u>Proposition 6</u> (Local Uniqueness)

<u>Let</u> g <u>be a proper function. Then the number of critical points of</u> $(P(y))$, <u>denoted</u> $\#(P(y))$, <u>is finite for any</u> $y \in Y$, <u>and it is a locally constant function on the open set</u> Y.

Proof

Let $(P(\bar{y}))$ be a proper Morse program, then $g^{-1}(\bar{y})$ is compact and each critical point of $(P(\bar{y}))$ is isolated. An isolated set in a compact set is finite, hence $\#(P(\bar{y}))$ is finite. Let $\#(P(\bar{y})) = k$ and let $\bar{x}^i$ and $\bar{\lambda}^i$ be respectively a critical point of $(P(\bar{y}))$ and its associated Lagrange multiplier $(i = 1,2,\cdots,k)$. By the implicit function theorem (Edwards [5] p. 417), for each $i = 1,\cdots,k$; there exist neighborhoods $W^i(\bar{y}) \subseteq Y$, $U^i(\bar{x}^i) \subseteq R^n$, $V^i(\bar{\lambda}^i) \subseteq R^m$, and C^1 functions $x^i(\cdot) : W^i \to U^i$, $\lambda^i(\cdot) : W^i \to V^i$ such that

$$(x^i(\bar{y}),\lambda^i(\bar{y})) = (\bar{x}^i,\bar{\lambda}^i)$$

$$F(x,\lambda,y) = 0 \iff (x,\lambda) = (x^i(y),\lambda^i(y)) \quad \text{on} \quad U^i \times V^i \times W^i \quad . \tag{4.6}$$

Now let us take a neighborhood W of $\bar{y}$ such that $W \subseteq \bigcap_{i=1}^{k} W^i$ and $x^1(W),\cdots,x^k(W)$ are pairwise disjoint.

Since $F(x^i(y),\lambda^i(y),y) = 0$ for $y \in W$, $x^i(y)$ is a critical point of $(P(y))$ for $i = 1,\cdots,k$. Since $x^1(W),\cdots,x^k(W)$ are pairwise disjoint, $x^1(y),\cdots,x^k(y)$ are k distinct critical points of $(P(y))$. Therefore we obtain $\#(P(y)) \geq k$ for $y \in W$.

Let us show that actually equality holds. Suppose, to the contrary, there exists $\{y^\ell\}$ such that $y^\ell \in W$, $y^\ell \to \bar{y}$ and $\#(P(y^\ell)) > k$. Then there exists $\{(x^\ell,\lambda^\ell)\}$ such that $F(x^\ell,\lambda^\ell,y^\ell) = 0$ and $x^\ell \notin \{x^1(y^\ell),\cdots,x^k(y^\ell)\}$. Note that x^ℓ is a critical point of $(P(y^\ell))$ with the associated Lagrange multiplier λ^ℓ.

Take $\varepsilon > 0$ so that a closed ε-ball $B_\varepsilon(\bar{y}) \subseteq W$, then there exists L such that $\ell > L \Rightarrow y^\ell \in B_\varepsilon(\bar{y})$. Since g is proper, $g^{-1}(B_\varepsilon(\bar{y}))$ is compact and $\ell > L \Rightarrow x^\ell \in g^{-1}(B_\varepsilon(\bar{y}))$. Then there exists a converging subsequence of $\{x^\ell\}_{\ell > L}$. For the notational convenience we assume $x^\ell \to x^*$ for some $x^* \in g^{-1}(B_\varepsilon(\bar{y}))$. Since $B_\varepsilon(\bar{y}) \subseteq Y$, $Dg(x)$ is of full rank for any $x \in g^{-1}(B_\varepsilon(\bar{y}))$. Therefore, $\{x^\ell\}_{\ell > L} \subseteq g^{-1}(B_\varepsilon(\bar{y}))$, $x^* \in g^{-1}(B_\varepsilon(\bar{y}))$, $x^\ell \to x^*$ imply $\lambda^\ell \to \lambda^*$ for some $\lambda^* \in R^m$ by (4.5). Then by the continuity of F, we obtain $F(x^*,\lambda^*,\bar{y}) = 0$. This implies $(x^*,\lambda^*) = (\bar{x}^j,\bar{\lambda}^j)$ for some $j \in \{1,\cdots,k\}$, hence for sufficiently large ℓ, we obtain

$$F(x^\ell,\lambda^\ell,y^\ell) = 0 \quad , \quad (x^\ell,\lambda^\ell,y^\ell) \in U^j \times V^j \times W^j$$

and

$$x^\ell \notin \{x^1(y^\ell),\cdots,x^k(y^\ell)\} \quad .$$

This contradicts to (4.6), so completes the proof.

Q.E.D.

Note that $\omega(y) = \min\limits_{1 \le i \le k} f(x^i(y))$ for $y \in W$, and hence we obtain

Corollary 7

If $(P(\bar{y}))$ is a proper Morse program, then in a neighborhood of $\bar{y}$, $\omega(\cdot)$ is the minimum of a finite number $(= \#(P(\bar{y}))$ of C^2 functions (as a result $\omega(\cdot)$ is a locally Lipschitz function on the open set Y). Moreover, if $(P(\bar{y}))$ has a unique global solution, then in a neighborhood of $\bar{y}$, $\omega(\cdot)$ is a C^2 function.

Proof

It is easily verified (e.g. Luenberger [11], 10.5) that

$$Df(x^i(y)) = -\lambda^i(y),\ D^2f(x^i(y)) = -D\lambda^i(y)\ (i = 1,\cdots,k) \quad ,$$

hence $f(x^i(\cdot))$ is in C^2. The fact that $\omega(\cdot)$ is locally Lipschitz on Y follows from Clarke [4].

Q.E.D.

Remark

The differentiability of the local optimum value function was shown by Fiacco/McCormick [6] and Fiacco [7], using the implicit function theorem. In our proper programs, we consider the global optimum value function $\omega(\cdot)$, and we will show in Theorem E(d), that the global optimum value function

$$\omega(u,v) : = \text{minimum}\{f(x) - u^T x \ \text{ subject to } \ g(x) = b + v\}$$

for $(P(u,v))$ is C^2 function of (u,v) on an open and dense set of $R^n \times R^m$, if $f \in C^2$ and $g \in C^{n-m+1}$.

Now let us make a few remarks on the open set Y.

Definition (Brown/Heal/Westhoff[3])

A program (P) is called regular if $F \pitchfork 0$.

The regularity is a generic property, namely if f and $g \in C^{m+2}$ (hence $F \in C^{m+1}$), then $F \pitchfork (u,v)$ for almost every $(u,v) \in R^n \times R^m$ by Sard's theorem, hence $(P(u,v))$ is regular for almost every $(u,v) \in R^n \times R^m$. If

$f,g \in C^{m+2}$ and $F \pitchfork 0$, then $F_y \pitchfork 0$ for almost every $y \in R^m$ by virtue of the parametric transversality theorem (Appendix (4)). Since $g \in C^{n-m+1}$, this implies $g \pitchfork y$ for almost every $y \in R^m$, summarizing the above we obtain

Proposition 8 (Brown/Heal/Westhoff [3])

(a) If $f,g \in C^{m+2}$ then $(P(u,v))$ is a regular program for almost every $(u,v) \in R^n \times R^m$.

(b) If (P) is regular, $f \in C^{m+2}$, $g \in C^{\max(m+2,n-m+1)}$ then Y has full measure in R^m.

Corollary 9 (cf. [3])

If (P) is a proper regular program, $f \in C^{m+2}$, and $g \in C^{\max(m+2,n-m+1)}$, then Y is open and dense in R^m.

Our final result in this section considers the differentiability of the global optimum value function $\omega(\cdot,\cdot)$ for $(P(u,v))$.

First let us define $\bar{F} : R^n \times R^m \times R^n \times R^m \to R^n \times R^m$ by

$$\bar{F}(x,\lambda,u,v) := (Df(x)^T - u + Dg(x)^T\lambda,\ g(x) - b - v)$$

and $\bar{F}_{(u,v)} : R^n \times R^m \to R^n \times R^m$ by

$$\bar{F}_{(u,v)}(x,\lambda) := F(x,\lambda,u,v) \quad .$$

Then we have

$$D\bar{F}_{(u,v)}(x,\lambda) = \begin{pmatrix} L(x) & Dg(x)^T \\ Dg(x) & 0 \end{pmatrix}$$

where $L(x) = D^2 f(x) + \sum_1^m \lambda_i D^2 g_i(x)$. Therefore following exactly the same argument as in Theorem B, we obtain

Lemma 10

If $g \pitchfork b + v$, *then* $f(x) - u^T x$ *is a Morse function on* $g^{-1}(b+v)$ *if and only if* $\bar{F}_{(u,v)} \pitchfork 0$.

Let us denote

$$Z := \{(u,v) \in R^n \times R^m \,|\, (P(u,v)) \text{ is a Morse program}\}$$

and

$$Z^! := \left\{ (u,v) \in R^n \times R^m \middle| \begin{array}{l} (P(u,v)) \text{ is a Morse program} \\ \text{having a unique global solution} \end{array} \right\}.$$

Then by Lemma 10 we have

$$Z = \{(u,v) \in R^n \times R^m \,|\, g \pitchfork b + v,\ \bar{F}_{(u,v)} \pitchfork 0\} .$$

Note that if $f \in C^2$ and $g \in C^{n-m+1}$, then Z has full measure in $R^n \times R^m$. Because $\{(u,v) \,|\, g \pitchfork b + v\}$ has full measure and $\{(u,v) \,|\, \bar{F}_{(u,v)} \pitchfork 0\}$ has full measure by Sard's theorem.

Following essentially the same argument as in Lemma 5 and Proposition 6, it is easy to prove the following

Theorem E

Suppose g *is a proper function. Then we have*

(a) Z *and* $Z^!$ *are open sets of* $R^n \times R^m$

(b) *the number of critical points of* $(P(u,v))$ *is finite for any* $(u,v) \in Z$, *and it is locally constant on* Z.

(c) $\omega(\cdot,\cdot)$ *is locally expressive as the minimum of a finite number of* C^2 *functions on* Z.

(d) *If* $f \in C^2$ *and* $g \in C^{n-m+1}$, *then* $\omega(\cdot,\cdot)$ *is a* C^2 *function on* $Z^!$ *which is open and dense in* $R^n \times R^m$.

5. Inequality Constraints: Definition of Morse Programs

Let us consider a program

(Q): $\quad\quad$ $\text{minimize}\{f(x) \text{ subject to } g(x) \leq b\}$

and a perturbation

$(Q(u,v))$: $\quad$ $\text{minimize}\{f(x) - u^T x \text{ subject to } g(x) \leq b + v\}$

where $f : R^n \to R^1$ and $g : R^n \to R^m$ are of class C^2, $u \in R^n$, $v \in R^m$ and $n \geq m$.

Let $I := \{1,2,\cdots,m\}$; $g_J(x) := (g_j(x))_{j\in J}$, $b_J := (b_j)_{j\in J}$, $J^c := I - J$ for all $J \subseteq I$. Let us denote

$$M_{J,i} := \{x \mid g_J(x) = b_J,\ g_i(x) \leq b_i\}$$

$$\partial M_{J,i} := \{x \mid g_J(x) = b_J,\ g_i(x) = b_i\}$$

for all $J \subseteq I$ and $i \in I$. For the notational convenience, we denote $M_J :=$ $M_{J,i}$ if $i \in J$ and $X_i := M_{J,i}$, $\partial X_i := \partial M_{J,i}$ if $J = \phi$. Note that if $i \in J$ then $M_{J,i} = \partial M_{J,i} = M_J$.

We will reduce the inequality case to some equalities and one inequality case.

<u>Definition</u>. A program (Q) is called a <u>Morse program</u> if (Q) satisfies

(M1) $\quad g_i \pitchfork b_i$ for all $i \in I$

(M2) $\quad g_J|_{X_i} \pitchfork b_J$ and $g_J|_{\partial X_i} \pitchfork b_J$ for all nonempty $J \subseteq I$ and $i \notin J$

(M3) $\quad f$ is a Morse function on $M_{J,i}$ and $\partial M_{J,i}$ for all $J \subseteq I$ and $i \in I$

(M4) $\quad f|_{M_{J,i}}$ has no critical points on $\partial M_{J,i}$ for all $J \subseteq I$ and $i \notin J$.

Remark

By (M1), $X_i = \{x | g_i(x) \leq b_i\}$ is n-dimensional manifold with boundary $\partial X_i = g_i^{-1}(b_i)$ (Appendix (10)). Then (M2) implies that $M_{J,i} = M_J \cap X_i$ is $(n - |J|)$-dimensional manifold with boundary $\partial M_{J,i} = M_J \cap \partial X_i = M_{J \cup \{i\}}$ (Appendix (11)). These manifolds of different dimensions cover the feasible region of (Q), and by (M3) we will restrict the objective function f on each manifold when we argue the optimality conditions of (Q). In the next section (Proposition 15), we will show that assuming (M1) and (M2), (M4) implies the strict complementary slackness condition.

Definition. x is a critical point of (Q) if $g(x) \leq b$ and x is a critical point of $M_{J(x)}$ where $J(x) := \{i | g_i(x) = b_i\}$.

The next theorem states the important properties of a Morse program, which is analogous to Theorem A.

Theorem F. If (Q) is a Morse program and x is a critical point of (Q) with $J = J(x)$, then

(a) $Dg_J(x)$ has full rank.

(b)[1] there exists a unique $\lambda \in R^m$ such that $Df(x)^T + Dg(x)^T\lambda = 0$ and $\lambda_i \neq 0$ iff $i \in J$.

(c) $L(x) = D^2f(x) + \sum_{i=1}^{m} \lambda_i D^2 g_i(x)$ induces an isomorphism on $T_x M_J$.

(d) On $T_x M_J$, $L(x)$ is positive definite iff x is a local minimum; negative definite iff x is a local maximum; indefinite iff x is a saddle point on M_J.

---

1) $\lambda_j > 0$ for all $j \in J$ if x is a local minimum, $\lambda_j < 0$ for all $j \in J$ if x is a local maximum (see Luenberger [11], 10.6).

Proof. (a) and (c) follow from (M1), (M2) and Lemma 3. (b) follows from Lemma 1 and (M4) (see Proposition 15). Since a local minimum (or maximum) point of (Q) is also a local minimum (or maximum) point of f on M_J, (d) follows from Theorem A(d).

Q.E.D.

6. Inequality Constraints: Generic Properties of Morse Programs

We will show that if $f \in C^2$ and $g \in C^n$, then for almost every fixed $v \in R^m$, $(Q(u,v))$ is a Morse program for almost every $u \in R^n$. First of all we consider the genericity of properties (M1) - (M3).

Lemma 11. If $g \in C^n$, then (M1) and (M2) hold for almost every $b \in R^m$.

Proof. We follow the proof of Spingarn/Rockafellar [17], Theorem 1. By Sard's theorem, if $g \in C^n$ then the set of critical values of g_i is of measure zero in $R^{|i|}$ for $i = 1,\ldots,m$. Then $T_i := \{b \in R^m | b_i$ is a critical value of $g_i\}$ is of measure zero in R^m $(i = 1,\ldots,m)$, because every $(m-|i|)$-dimensional horizontal slice of T_i is of measure zero as a subset of $R^{|i|}$, hence T_i itself must have measure zero by Fubini's theorem (Appendix (3)). Hence $T = \bigcup_{i=1}^{m} T_i$ has measure zero. By Sard's theorem with boundary (Appendix (2)), if $g \in C^n$ and $b \in R^m - T$ (hence X_i is n-dimensional manifold with boundary ∂X_i), then for any nonempty $J \subseteq I$ and any $i \notin J$, the set of critical values of $g_J|_{X_i}$ or $g_J|_{\partial X_i}$ is of measure zero in $R^{|J|}$. So, again, by Fubini's theorem (Appendix (3)),

$$S_{J,i} := \left\{ b \in R^m - T \,\middle|\, \begin{array}{l} b_J \text{ is a critical value of} \\ g_J|_{X_i} \text{ or } g_J|_{\partial X_i} \end{array} \right\}$$

has measure zero in R^m for all nonempty $J \subseteq I$ and all $i \notin J$. Then $S := \bigcup_{J,i} S_{J,i}$ has measure zero in R^m.

Q.E.D.

Let us take $b \in S = \{b \in R^m |\ b$ satisfies (M1) and (M2)$\}$. Then for any $J \subseteq I$ and $i \in I$, $f(x) - u^T x$ is a Morse function on $M_{J,i}$ and $\partial M_{J,i}$ for almost every $u \in R^n$ by Appendix (6). Therefore we obtain

Corollary 12. *If* $f \in C^2$ *and* $g \in C^n$, *then for almost every* $v \in R^m$, $(Q(u,v))$ *satisfies* (M1), (M2) *and* (M3) *for almost every* $u \in R^n$.

The genericity of the strict complementary slackness condition is much more complicated and we need some preliminary results.

Proposition 13

Let M *be an* m-*dimensional* C^r *manifold in* R^n *with nonempty boundary* ∂M *and let* $f : M \to R^1$ *be a* C^r *map. Then for almost every* $u \in R^n$, C^r *map* $f(x) - u^T x$ *defined on* M *has no critical point on* ∂M.

Proof[1] $x \in M$ is a critical point of $f(x) - u^T x$ iff $Df(x)^T - u \perp T_x M$ iff $u \in Df(x)^T + T_x M^{\perp}$ ($T_x M^{\perp}$ is the orthogonal complement of $T_x M$ in R^n). Let $E := \{(x,u) \in \partial M \times R^n | u \in Df(x)^T + T_x M^{\perp}\}$. Then E is $(n-1)$-dimensional C^{r-1} submanifold of $\partial M \times R^n$. Let us prove this fact. For any given $\bar{x} \in \partial M$ there exists an open set $\tilde{U}$ of R^n and a submersion $g : \tilde{U} \to R^{n-m}$ such that $U := M \cap \tilde{U} = g^{-1}(0)$ and $\bar{x} \in \partial U := \partial M \cap \tilde{U}$ (Appendix (7)). Let $\Phi : \partial U \times R^{n-m} \to \partial M \times R^n$ be $\Phi(x,y) := (x, Df(x)^T + Dg(x)^T y)$; (ϕ, V) be a local parametrization of ∂M at $\bar{x}$ such that $\bar{x} = \phi(\bar{v})$, $\bar{v} \in V \subseteq R^{m-1}$, and $\phi(v) \subseteq \partial U$; and $\psi : V \times R^{n-m} \to \partial M \times R^n$ be $\psi(v,y) = \Phi(\phi(v), y)$. Then $\psi \in C^{r-1}$ and for any $(v,y) \in V \times R^{n-m}$ we have

$$D\psi(v,y) = \begin{bmatrix} D\phi(v) & 0 \\ * & Dg(\phi(v))^T \end{bmatrix} .$$

---

1) This line of argument was suggested by W. G. Dwyer.

Since $D\phi(v)$ and $Dg(\phi(v))^T$ is 1-1, $D\psi(v,y)$ is 1-1 (i.e. ψ is an immersion) hence by the definition Φ is an immersion. Let $E(\partial U) = \{(x,u) \in \partial U \times R^n | u \in Df(x)^T + T_x M^\perp\} \subseteq \partial M \times R^n$, then since $T_x M^\perp = \mathrm{Im}\ Dg(x)^T$ $\Phi : \partial U \times R^{n-m} \to E(\partial U)$ is bijective and proper, hence Φ is an embedding of $\partial U \times R^{n-m}$ into $\partial M \times R^n$. Consequently $E(\partial U)$ is a C^{r-1} manifold (Appendix (8)) parametrized by Φ, with dimension $= \dim \partial U + n - m = m - 1 + n - m = n - 1$. Since every point of E has such a neighborhood, E is a C^{r-1} manifold. (cf. Guillemin/Pollack [8], normal bundle on page 71.) Let $\pi : \partial M \times R^n \to R^n$ be a projection map. Then since E is $(n-1)$-dimensional, $\pi(E) \subseteq R^n$ has measure zero in R^n (Appendix (9)).

Since $\pi(E) = \{u \in R^n | u \in Df(x)^T + T_x M^\perp$ for some $x \in \partial M\}$ has measure zero, for almost every $u \in R^n$ (i.e., $u \notin \pi(E)$) every $x \in \partial M$ is not a critical point of $f(x) - u^T x$ on M. This completes the proof.

Q.E.D.

<u>Lemma 14. Let</u> $g : R^n \to R^m$, $h : R^n \to R^1$; $b \in R^m$, $c \in R^1$; $n \geq m + 1$; $X := \{x | h(x) \leq c\}$, <u>and</u> $\partial X = h^{-1}(c)$. <u>Suppose</u> $h \pitchfork c$, <u>then we have</u>

(a) <u>If</u> $g_{|X} \pitchfork b$, $g_{|\partial X} \pitchfork b$ <u>then</u> $M := g^{-1}(b) \cap X$ <u>is</u> $(n-m)$-<u>dimensional manifold with boundary</u> $\partial M := g^{-1}(b) \cap h^{-1}(c)$, <u>and</u> $T_x \partial M = \mathrm{Ker}\ Dg(x) \cap \mathrm{Ker}\ Dh(x)$.

(b) $g_{|\partial X} \pitchfork b$ <u>iff</u> $(g,h) \pitchfork (b,c)$.

<u>Proof</u>. Since $h \pitchfork c$, X is n-dimensional manifold with boundary $\partial X = h^{-1}(c)$ by Appendix (10).

(a): By Appendix (11), M is $(n-m)$-dimensional manifold with boundary $\partial M = g^{-1}(b) \cap h^{-1}(c)$. $g_{|\partial X} \pitchfork b$ and $T_x \partial X = \mathrm{Ker}\ Dh(X)$ imply $T_x \partial M = \mathrm{Ker}\ Dg(x) \cap \mathrm{Ker}\ Dh(x)$ by Appendix (12).

(b): (only if): Since $\dim T_x \partial M = n-m-1$, $\begin{pmatrix} Dg(x) \\ Dh(x) \end{pmatrix} : R^n \to R^m \times R^1$ is onto for any $x \in \partial M$, hence $(g,h) \pitchfork (b,c)$.

(if): Let $x \in (g_{|\partial X})^{-1}(b)$ i.e. $x \in (g,h)^{-1}(b,c)$. We will show that $Dg(x)_{|T_x \partial X} : T_x \partial X \to R^m$ is onto. Since $T_x \partial X =$ Ker $Dh(x)$, Ker$\left(Dg(x)_{|T_x \partial X}\right) =$ Ker $Dg(x) \cap$ Ker $Dh(x)$. However $\dim(\text{Ker } Dg(x) \cap \text{Ker } Dh(x)) = n-m-1$ because $(g,h) \pitchfork (b,c)$, and since $\dim T_x \partial X = n-1$, we obtain $Dg(x)_{|T_x \partial X}$ is onto. Hence $g_{|\partial X} \pitchfork b$.

Q.E.D.

<u>Theorem G</u>. (Strict Complementary Slackness)

<u>Consider a nonlinear programming problem</u>

<u>minimize</u> $\{f(x)$ <u>subject to</u> $g(x) = b,\ h(x) \leq c\}$

<u>where</u> $f : R^n \to R^1$, $g : R^n \to R^m$, $h : R^n \to R^1$; $f,g,h \in C^1$; $n \geq m+1$.

<u>Let</u> $X := \{x | h(x) \leq c\}$, $\partial X := h^{-1}(c)$, $M := g^{-1}(b) \cap X$, <u>and</u> $\partial M := g^{-1}(b) \cap \partial X$. <u>Suppose</u> $h \pitchfork c$, $g_{|X} \pitchfork b$ <u>and</u> $g_{|\partial X} \pitchfork b$. <u>Let</u> x^* <u>be a critical point of</u> $f_{|\partial M}$. <u>Then the Lagrange multiplier</u> μ^* <u>associated with the constraint</u> $h(x) \leq c$ <u>is nonzero if and only if</u> x^* <u>is not a critical point of</u> $f_{|M}$.

<u>Proof</u>. By Lemma 14(a), M is $(n-m)$-dimensional manifold with boundary ∂M. Note that x^* is a critical point of $f_{|\partial M}$ if and only if $Df(x^*) \in T_{x^*} \partial M^{\perp}$. We have $T_{x^*} \partial M^{\perp} = \text{Im}(Dg(x^*)^T, Dh(x^*)^T)$ because $(g,h) \pitchfork (b,c)$ by Lemma 14(b), hence there exists a unique $(\lambda^*, \mu^*) \in R^m \times R^1$ such that

$$(6.1) \qquad Df(x^*)^T + Dg(x^*)^T \lambda^* + Dh(x^*)^T \mu^* = 0 \quad .$$

If x^* is not a critical point of $f_{|M}$, then $Df(x^*)^T \notin T_{x^*}M^\perp = \operatorname{Im} Dg(x^*)^T$ and hence in (6.1), $\mu^* \neq 0$. On the other hand if x^* is a critical point of $f_{|M}$, then by the uniqueness of (λ^*, μ^*) in (6.1), $\mu^* = 0$.

Q.E.D.

This theorem provides a geometric interpretation of the strict complementary slackness, namely the degenerate Lagrange multiplier occurs if and only if the critical point of f on the boundary of the manifold M is also a critcal point of $f_{|M}$. Let us illustrate this fact by an example.

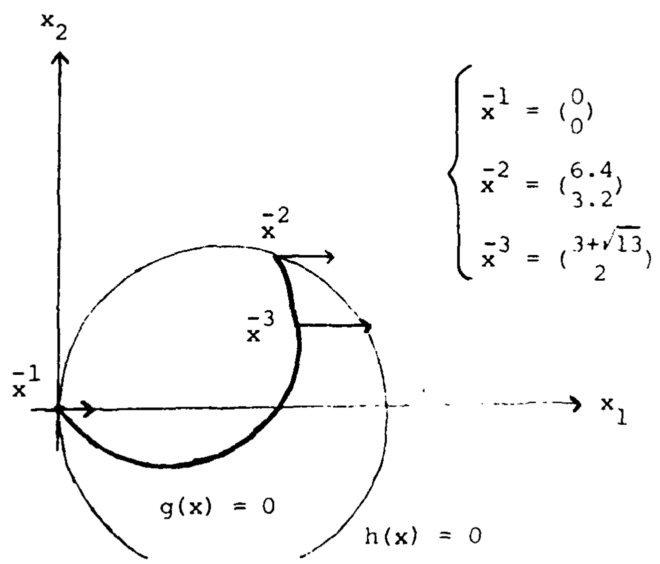
Example (Avriel [2], Example 3.1.4)

Consider the following program:

$$\text{minimize } f(x) = x_1$$

$$\text{subject to } g(x) = (x_1-3)^2 + (x_2-2)^2 - 13 = 0$$

$$h(x) = (x_1-4)^2 + x_2^2 - 16 \leq 0$$



It is easily verified that $g \pitchfork 0$, $h \pitchfork 0$. $g_{|h^{-1}(0)} \pitchfork 0$ is obvious because $\{x|g(x) = 0\}$ meets $\{x|h(x) = 0\}$ <u>transversally</u>. Hence all assumptions in Theorem G are satisfied.

Let $M := \{x|g(x) = 0,\ h(x) \leq 0\}$, then $\partial M := \{\bar{x}^1, \bar{x}^2\}$. Since every point is a critical point of any function defined on 0-dimensional manifold, $\bar{x}^1$ and $\bar{x}^2$ are critical points of $f_{|\partial M}$. Since $Df(x) = (1,0)$ for any $x \in R^n$, $f_{|M}$ has only one critical point $\bar{x}^3$ which is not an element of ∂M. Therefore by Theorem G, the associated Lagrange multiplier of the constraint $h(x) \leq 0$, $\bar{\mu}^1$ (or $\bar{\mu}^2$) at $\bar{x}^1$ (or $\bar{x}^2$) is nonzero.

Now we can show that (M4) implies the strict complementary slackness condition.

<u>Proposition 15</u>

<u>Suppose</u> (Q) <u>satisfies</u> (M1), (M2) <u>and</u> (M4). <u>Then every Lagrange multiplier associated with an active constraint is nonzero</u> (i.e. <u>strict complementary slackness condition holds</u>).

<u>Proof</u>

Let $\bar{x}$ be a critical point of (Q) such that $J = J(\bar{x})$. Then by (M1) and (M2), there exists a unique $\bar{\lambda} \in R^m$ such that $\bar{\lambda}_{J^c} = 0$ and

$$Df(\bar{x})^T + \sum_{j \in J} \bar{\lambda}_j Dg_j(\bar{x}) = 0 \quad .$$

To show $\bar{\lambda}_j \neq 0$ for all $j \in J$, pick any $j \in J$. By (M1) and (M2), we have

$$g_j \pitchfork b_j,\ g_{J-\{j\}}{}_{|X_j} \pitchfork b_{J-\{j\}},\ g_{J-\{j\}}{}_{|\partial X_j} \pitchfork b_{J-\{j\}} \quad .$$

Since $\bar{x}$ is a critical point of (Q), this means $\bar{x}$ is a critcal point of $f_{|M_J}$ and $M_J = \partial M_{J-\{j\},j}$. By (M4), $\bar{x}$ is not a critical point of $f_{|M_{J-\{j\},j}}$. Hence by Theorem G, the Lagrange multiplier $\bar{\lambda}_j$ associated with the constraint $g_j(x) \leq b_j$ is nonzero.

Q.E.D.

By Proposition 13, the genericity of (M4) is obtained and hence we have

Theorem H (cf. Spingarn/Rockafellar [17], Corollary)

If $f \in C^2$ and $g \in C^n$, then for almost every fixed $v \in R^m$, $(Q(u,v))$ is a Morse program having at most one global solution for almost every $u \in R^n$.

Proof

By Araujo/Mascolell ([1]), we have also

Fix any $v \in R^m$, then for almost every $u \in R^n$ $(Q(u,v))$ has at most one global solution.

Use Corollary 12, and Proposition 13.

Q.E.D.

Acknowledgement

I wish to thank my thesis advisor, Professor Donald J. Brown of Yale University, for his many helpful suggestions concerning this material. I also wish to thank the referee for his helpful remarks.

Appendix

(Guillemin/Pollack [8], Hirsch [10])

(1) Morse lemma

Let $p \in M$ be a nondegenerate critical point of $f : M \to R^1$. Then there is a local coordinate system $(x_1, \ldots, x_m)$ in a neighborhood U of p such that

$$f = f(p) - x_1^2 - \ldots - x_\lambda^2 + x_{\lambda+1}^2 + \ldots + x_m^2$$

for some $0 < \lambda < m$.

(2) Sard's theorem (with boundary)

Let $f : X \to Y$ be a C^γ map of a C^γ manifold X with boundary ∂X into a boundaryless C^γ manifold Y. Then almost every $y \in Y$ is a regular value of both $f : X \to Y$ and $f|_{\partial X} : \partial X \to Y$ if $\gamma > \max(0, \dim X - \dim Y)$.

(3) Fubini's theorem

Let $A \subseteq R^n \times R^m$ be a measurable set such that for almost every $v \in R^m$, $A_v = \{u \in R^n | (u,v) \in A\}$ has measure zero in R^n. Then A has measure zero in $R^n \times R^m$.

(4) Parametric transversality theorem

Let $F : X \times V \to Y$ be a C^γ map of C^γ manifolds and A be any C^γ submanifold of Y. If $F \pitchfork A$ and $\gamma > \max(0, \dim X - \dim Y)$ then $F_v \pitchfork A$ for almost every $v \in V$ where $F_v(x) = F(x,v)$ for $x \in X$.

(5) Let $f : X \to Y$ be a C^γ map such that $f \pitchfork Z$ for a C^γ submanifold Z of Y, then $f^{-1}(Z)$ is a C^γ submanifold of X and $\dim f^{-1}(Z) =$

$\dim X - \dim Y + \dim Z$. As a special case if $f \pitchfork y$ for some $y \in Y$, then $f^{-1}(y)$ is a C^{γ} submanifold of X and $\dim f^{-1}(y) = \dim X - \dim Y$.

(6) Let $f : X \to R^1$ be a C^2 map of a C^2 manifold X in R^n. Then for almost every $u \in R^n$ the function $f(x) - u^T x$ is a Morse function on X.

(7) Let $X \subseteq R^n$ be m-dimensional manifold with boundary ∂X. Then for each point $x \in \partial X$, there exists an open set $\tilde{U}$ of R^n and a submersion $g : \tilde{U} \to R^{n-m}$ such that $U = X \cap \tilde{U} = g^{-1}(0)$ and $x \in \partial U = \partial X \cap \tilde{U}$.

(8) An embedding $f : X \to Y$ maps X differomorphically onto a submanifold of Y.

(9) Let X, Y be manifolds with $\dim X < \dim Y$. If $f : X \to Y$ is a C^1 map then $f(X)$ has measure zero in Y.

(10) Let $f : X \to R^1$ be a C^{γ} map such that $f \pitchfork c$ for some $c \in R^1$. Then $\{x | f(x) \leq c\}$ is a C^{γ} submanifold of X with boundary $f^{-1}(c)$.

(11) Let $f : X \to Y$ be a C^{γ} map of a C^{γ} manifold X with boundary ∂X onto a boundaryless C^{γ} manifold Y. If $f \pitchfork Z$, $f|_{\partial X} \pitchfork Z$ for a boundaryless submanifold Z of Y, then $f^{-1}(Z)$ is a C^{γ} submanifold of X with boundary $\partial f^{-1}(Z) = f^{-1}(Z) \cap \partial X$ and $\dim f^{-1}(Z) = \dim X - \dim Y + \dim Z$.

(12) Let X, Z be submanifolds of Y such that $X \pitchfork Z$. Then $X \cap Z$ is again a submanifold of Y, $\dim(X \cap Z) = \dim X + \dim Z - \dim Y$ and $T_x(X \cap Z) = T_x X \cap T_x Z$ for any $x \in X \cap Z$.

More generally, let $f : X \to Y$ be a map transversal to a submanifold Z in Y. Then $W = f^{-1}(Z)$ is a submanifold of X and $T_x W = \operatorname{Ker} Df_x$ where $Df_x : T_x X \to T_{f(x)} Y$.

REFERENCES

1. Araujo, A. and A. Mas-Colell, "Notes on the Smoothing of Aggregate Demand," J. of Math. Econ., 5 (1978), 113-127.

2. Avriel, M., Nonlinear Programming, Analysis and Methods, Prentice-Hall, (1976).

3. Brown, D. J., G. Heal and F. Westhoff, "Regular Nonlinear Programs," Preprint.

4. Clarke, F. H., "A New Approach to Lagrange Multipliers," Math. of Oper. Res., 1 (1976), 165-174.

5. Edwards, C. H. Jr., Advanced Calculus of Several Variables, Academic Press, New York, 1973.

6. Fiacco, A. V. and G. P. McCormick, Nonlinear Programming: Sequential Unconstrained Minimization Techniques, John Wiley, New York, 1968.

7. Fiacco, A. V., "Sensitivity Analysis for Nonlinear Programming Using Penalty Methods," Math. Prog. 10 (1976), 287-311.

8. Guillemin, V. and A. Polack, Differential Topology, Prentice-Hall, 1974.

9. Hestenes, M. R., Optimization Theory: The Finite Dimensional Case, John Wiley, 1975.

10. Hirsch, M. R., Differential Topology, Springer-Verlag, 1976.

11. Luenberger, D. G., Introduction to Linear and Nonlinear Programming, Addison Wesley, 1973.

12. Mangasarian, O. L., Nonlinear Programming, McGraw-Hill, New York, 1969.

13. Rockafellar, R. T., Convex Analysis, Princeton University Press, Princeton, NJ, 1970.

14. Spingarn, J. E., "Generic Conditions for Optimality in Constrained Minimization Problems," Dissertation, Dept. of Mathematics, University of Washington (1977).

15. Spingarn, J. E., "Fixed and Variable Constraints in Sensitivity Analysis," SIAM J. Control and Optimization, 18, 3 (1980), 297-310.

16. Spingarn, J. E., "On Optimality Conditions for Structured Families on Nonlinear Programming Problems," Preprint.

17. Spingarn, J. E. and R. T. Rockafellar, "The Generic Nature of Optimality Conditions in Nonlinear Programming," Math. of Oper. Res., 4 (1979), 425-430.

18. Tanabe, K., "An Algorithm for Constrained Maximization in Nonlinear Programming," J. of Oper. Res. of Japan, 17 (1974), 184-201.

19. Tanabe, K., "A Geometric Method in Nonlinear Programming," J. of Optimization Theory and Applications, 30 (1980), 181-210.

OF/jvs

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER #2169	2. GOVT ACCESSION NO. AD-A099357	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MORSE PROGRAMS: A TOPOLOGICAL APPROACH TO SMOOTH CONSTRAINED OPTIMIZATION		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Okitsugu Fujiwara		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041 ENG-78-25182 SOC-77-03277
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 5 - Operations Research
11. CONTROLLING OFFICE NAME AND ADDRESS See Item 18 below		12. REPORT DATE January 1981 13. NUMBER OF PAGES 34
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

and

National Science Foundation
Washington, D.C. 20550

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Morse functions, Nonlinear programming, Second order sufficiency conditions, Strict complementary slackness, Sensitivity analysis.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We consider nonlinear constrained optimization problems whose objective and constraint functions are sufficiently smooth. No convexity is assumed.

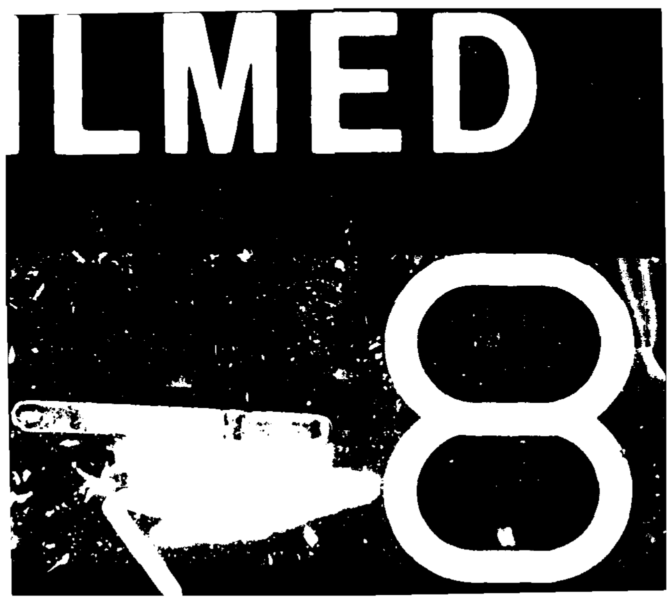
Our basic tools are from differential topology. We show that these problems can be reduced to the study of minimizing a Morse function on a manifold with boundary and we give the geometrical meaning to the first order conditions, the second order sufficiency conditions, and strict complementary slackness condition.

Our main concerns are the second order sufficiency conditions, sensitivity analysis, generic properties of smooth nonlinear programs, global duality, local uniqueness, and strict complementary slackness.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)


ILMED
8